Animal Behaviour 95 (2014) 89–99

Contents lists available at ScienceDirect

Animal Behaviour

journal homepage: www.elsevier.com/locate/anbehav

Vocal indicators of body size in men and women: a meta-analysis

Katarzyna Pisanski [a], Paul J. Fraccaro [a], Cara C. Tigue [a], Jillian J. M. O'Connor [a], Susanne Röder [b, c, d], Paul W. Andrews [a], Bernhard Fink [b, c], Lisa M. DeBruine [e], Benedict C. Jones [e], David R. Feinberg [a, *]

[a] Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, ON, Canada
[b] Institute of Psychology, University of Göttingen, Göttingen, Germany
[c] Courant Research Centre Evolution of Social Behavior, University of Göttingen, Göttingen, Germany
[d] Department of General Psychology and Methodology, University of Bamberg, Bamberg, Germany
[e] Institute of Neuroscience and Psychology, University of Glasgow, Glasgow, U.K.

ARTICLE INFO

Article history:
Received 31 January 2014
Initial acceptance 16 April 2014
Final acceptance 4 June 2014
Published online
MS. number: A14-00100R

Keywords:
bioacoustics
formant
fundamental frequency
height
human
sexual selection
source filter
vocal tract
voice
weight

Animals often use acoustical cues, such as formant frequencies, to assess the size of potential mates and rivals. Reliable vocal cues to size may be under sexual selection. In most mammals and many other vertebrates, formants scale with vocal tract length allometrically and predict variation in size more reliably than fundamental frequency or pitch (F0). In humans, however, it is unclear from previous work how well voice parameters predict body size independently of age and sex. We conducted a meta-analysis to establish the strength of various voice–size relationships in adult men and women. We computed mean weighted correlations from 295 coefficients derived from 39 independent samples across five continents, including several novel and large cross-cultural samples from previously unpublished data. Where possible, we controlled for sample size, sample sex, mean age, geographical location, study year, speech type and measurement method, and ruled out publication bias. Eleven of 12 formant-based vocal tract length (VTL) estimates predicted men's and women's heights and weights significantly better than did F0. Individual VTL estimates explained up to 10% of the variance in height and weight, whereas F0 explained less than 2% and correlated only weakly with size within sexes. Statistically reliable size estimates from F0 required large samples of at least 618 men and 2140 women, whereas formant-based size estimates required samples of at least 99 men and 164 women. The strength of voice–size relationships varied by sample size, and in some cases sex, but was largely unaffected by other demographic and methodological variables. We confirm here that, analogous to many other vertebrates, formants provide the most reliable vocal cue to size in humans. This finding has important implications for honest signalling theory and the capacity for human listeners to estimate size from the voice.

Among most terrestrial mammals, including humans, the voice is produced by the larynx and subsequently filtered by the supralaryngeal vocal tract (henceforth, vocal tract; Titze, 1994). The vocal folds within the larynx vibrate to produce the fundamental frequency (F0) and corresponding harmonics that are perceived as voice pitch, whereas formants are the resonant frequencies of the vocal tract. Because of relatively minimal feedback of vocal tract energy on vocal fold vibration, the source-filter model of speech production treats F0 and formants as anatomically and functionally independent (Fant, 1960; Titze, 1994).

Source-filter theory was originally developed by speech scientists (Fant, 1960; Singh & Singh, 1976; Titze, 1994), but has since been applied to the study of nonhuman vocalizations (see, e.g. Fitch & Hauser, 1995, 2003; Ohala, 1983; Owren & Bernacki, 1998; Owren, Seyfarth, & Cheney, 1997; Rendall, Owren, & Rodman, 1998; Sommers, Moody, Prosen, & Stebbins, 1992; Taylor & Reby, 2010). Research has confirmed that F0 and formants are decoupled in most vertebrates by showing that changes in F0 and formants do not covary in heliox, a mixture of helium and oxygen that transmits sound twice as fast as does air. In a coupled vocal system, heliox causes both F0 and formants to shift upward, whereas in a decoupled system, only formants shift upward due to a shortened transit time of sound waves traveling up the vocal tract (Hess et al., 2006). Decoupling has been demonstrated in several species of birds, anurans, bats and many mammalian species including humans (see Fitch & Hauser, 2003).

* Correspondence: D. R. Feinberg, Department of Psychology, Neuroscience & Behaviour, McMaster University, 1280 Main Street West, Hamilton, ON L8S 4K1, Canada.
E-mail address: feinberg@mcmaster.ca (D. R. Feinberg).

http://dx.doi.org/10.1016/j.anbehav.2014.06.011

Fundamental and formant frequencies can provide reliable affective and inferential information (e.g. Morton, 1977; Rendall et al., 1998) as well as reliable cues to sex and age. In humans, for instance, both F0 and formants are typically lower among men than women (Titze, 1989) and lower among adults than prepubescent children (Hillenbrand, Getty, Clark, & Wheeler, 1995; Peterson & Barney, 1952). At puberty, testosterone thickens and lengthens boys' vocal folds, causing F0 to drop (Harries, Hawkins, Hacking, & Hughes, 1998), and directly affects F0 throughout adulthood (Abitbol, Abitbol, & Abitbol, 1999; Dabbs & Mallinger, 1999; Damrose, 2009). Testosterone may also contribute to the sexual dimorphism in formant frequencies between men and women (Bruckert, Liénard, Lacroix, Kreutzer, & Leboucher, 2006).

In addition to indicating sex and age, F0 and formants are the two key acoustic parameters that have traditionally been investigated as potential vocal indicators of body size in humans and in mammals more generally (for reviews see: González, 2006; Taylor & Reby, 2010). Here, work on formant production in animals, particularly nonhuman animals, has produced a number of novel and testable hypotheses. Of particular interest to the present study, Fitch (1994, 1997, 2000) proposed that formants could reliably indicate body size in most vertebrates because the vocal tract is constrained by skeletal structures related to body size, which in turn imposes a constraint on resonances of the vocal tract (i.e. formants). Longer vocal tracts are predicted to produce lower and more closely spaced formants. However, because thicker and longer vocal folds vibrate at lower frequencies, it is also possible that, in some species, larger individuals with larger larynges may produce lower F0 than smaller individuals (Fitch & Hauser, 2003; Morton, 1977; Titze, 1994). Moreover, because F0 and formants are typically decoupled in vertebrates, larger individuals with correspondingly longer vocal tracts may produce lower formants regardless of F0 and vice versa (Fitch, 2000; Fitch & Giedd, 1999; Fitch & Hauser, 2003).

Since the first empirical study less than 20 years ago (Fitch, 1997), there has been a surge of research testing the relative roles of F0 and formants as honest indicators of size in a wide range of species. This research has generally confirmed that both F0 and formants independently predict variation in body size among individuals of different species, breeds or clades (primates: Hauser, 1993; dogs: Riede & Fitch, 1999; Taylor, Reby, & McComb, 2008; anurans: Gingras, Boeckle, Herbst, & Fitch, 2013; for reviews see: Ey, Pfefferle, & Fischer, 2007; Fitch & Hauser, 2003; Taylor & Reby, 2010) and between males and females of the same species (e.g. in humans and nonhuman primates: Fitch & Giedd, 1999; González, 2006; Rendall, Kollias, Ney, & Lloyd, 2005; Pfefferle & Fischer, 2006). Studies of a number of mammalian species have found, however, that formants are a better predictor of size within sexes than is F0. This pattern of results has been observed in studies of rhesus macaques, *Macaca mulatta* (Fitch, 1997), Japanese macaques, *Macaca fuscata* (Masataka, 1994), colobus monkeys, *Colobus satanas* (Harris, Fitch, Goldstein, & Fashing, 2006), red deer, *Cervus elaphus* (Reby & McComb, 2003), fallow deer, *Dama dama* (Vannoni & McElligott, 2008), koalas, *Phascolarctos cinereus* (Charlton et al., 2011), elephant seals, *Mirounga leonina* (Sanvito, Galimberti, & Miller, 2007), and dogs, *Canis familiaris* (Plotsky, Rendall, Riede, & Chase, 2013; Riede & Fitch, 1999). It is of potential interest that both F0 and formants appear to predict the size of male but not female giant pandas, *Ailuropoda melanoleuca* (Charlton, Zhihe, & Snyder, 2009), whereas F0 is a better predictor of the size of female hamadryas baboons, *Papio hamadryas*, than are formants (Pfefferle & Fischer, 2006).

It is unclear whether formants predict body size within sexes more reliably than does F0 in humans. A large proportion of the variation in F0 and formants among humans can be attributed to pubertal expression of sex hormones (Abitbol et al., 1999; Harries et al., 1998) and to differences in body size among men, women and children (Fitch & Giedd, 1999; Smith & Patterson, 2005; Turner, Walters, Monaghan, & Patterson, 2009). However, when investigating voice–size relationships within age–sex classes, it is not clear from previous work whether any voice parameter reliably predicts variation in human body size. At the within-sex level, adult body size showed no significant physical relationship with F0 in more than 80% of correlations reported in published studies and no significant relationship with formants in more than 50% of correlations reported in published studies. The strength and direction of reported correlations ranges widely (range of r for F0 estimates of size $= -0.71$ to $+0.30$; range of r for formant estimates of size $= -0.58$ and $+0.32$).

In addition to controlling for sex and age, a number of demographic and methodological factors may contribute to the variation in reported relationships between the voice and physical body size across studies. These factors may include the size and geographical location of the sample, the length and content of recorded speech materials, or the equipment and techniques used to measure the voice and body size. Proper adjustment of software settings when measuring vocal parameters, particularly formants, as well as the physical properties of the vocal signal may also affect the strength of reported voice–size relationships (Fitch & Fritz, 2006). Finally, the robustness of formant–size relationships is further complicated by differences in the measures used to relate formant structure to vocal tract length or body size across studies. Traditionally, studies examined size in relation to individual formants (i.e. first to fourth formant, F1 to F4: González, 2004; Greisbach, 1999; Rendall et al., 2005). More recent work has utilized amalgamated measures of formant structure including mean formant frequency (Fn; Pisanski & Rendall, 2011), geometric mean formant frequency (MFF; Smith & Patterson, 2005), formant dispersion (D_f; Fitch, 1997), formant position (P_f; Puts, Apicella, & Cardenas, 2012), formant spacing (ΔF; Reby & McComb, 2003) and apparent vocal tract length derived from mean formants (VTL(F_i); adapted from Fitch, 1997) or from formant spacing (VTL(ΔF); Reby & McComb, 2003; see equations A1–A7 in the Appendix). Moreover, a variation of the formant-pattern latent variable model uses confirmatory factor analysis (CFA) to relate factor scores of formants to vocal tract length and height within sexes (Turner et al., 2009). These formant-based measures will henceforth be referred to as VTL (vocal tract length) estimates. It is unknown which of these 12 VTL estimates most reliably predicts human body size.

The present study was designed to establish the strength of various relationships between the voice and body size in men and women at the population level. Until now, the nature of these relationships in humans has been unclear and has been a source of extensive and prolonged debate among researchers. We present the results of a meta-analysis of 295 voice–size correlations derived from 39 human adult samples, including several novel and large cross-cultural samples. The 39 samples derive from North and South America, Europe, Asia and Africa. The goals of the meta-analysis were to (1) determine the degree to which F0 and formants predict the height and weight of either sex, (2) test the relative reliability of 12 VTL estimates in predicting size, (3) assess the effects of age and a number of other demographic or methodological factors on the strength of voice–size relationships, (4) determine the minimum sample sizes required in future studies to obtain reliable estimates of size from the voice, (5) evaluate the effect of sample size and evidence for publication bias in our sample and (6) compute population-level averages of voice and body size parameters.

Human vocal cues may be under sexual selection (Puts, Jones, & DeBruine, 2012) and may honestly indicate mate quality (Maynard